

Topics in Computational Biology

Iceland 2.6.9-6.6.9

This is a one-week course taught by Jotun Hein and Rune Lyngsø. It is aimed at researchers/students with a strong quantitative background (mathematics, physics, statistics, computer science,...) that wants an overview of computational biology with a wish to do research in the area. Each day is devoted to a topic.

The students will be doing projects in groups. All projects in the course will be one of three kinds:

- *Reading projects*, where the students will have to read up upon a topic and present and discuss it. Examples of such topics could be Population Genomics: 1000 genomes, Integrative Genomics: Basic data types, Comparative Genomics: Signals, Comparative Biology: Networks, Integrative Data Analysis – Mapping, Metabolomics, Proteomics, Transcriptomics, Genomic Dark Matter, LUCA – Last Universal Common Ancestor, Somatic Cell Genealogies,... Examples of projects can be found at <http://www.stats.ox.ac.uk/genome/readingprojects>.
- *Analysis projects*, where all students will analyze a data set from a series of angles: association mapping, genome annotation, regulatory signals, signatures of selection, ... Typically it will focus on a gene of special interest due to recent publications and would typically relate to some disease. Examples of projects can be found at <http://www.stats.ox.ac.uk/genome/analysisprojects>.
- *Project discussions/criticism*. The projects will describe an idea for possible research. This has to be elaborated upon. Examples of projects can be found at <http://www.stats.ox.ac.uk/genome/projects>.

Preparation of the project discussion will take 4-8 hours. It is recommended that the students read the following before the course starts:

Yang, Z. (2009) Computational Molecular Evolution chapter 1 OUP
Hein, Schierup and Wiuf (2005) Gene Genealogies, Variation and Evolution chapter 1 + 5.
Davies, Rafnar, Hellenthal and Hein (2009) Integrative Genomics and Functional Explanation
Lyngsø (2009) RNA Structure Prediction Notes

Beyond this the course should be self-contained. The project presentations are expected to last 40-90 minutes.

Lectures, Practicals, Exercises and Project Preparation each lasts 90 minutes.

Monday June 2nd: Molecular Evolution

9.00-10.15 Lecture: Models of Sequence Evolution and Statistical Alignment
10.30-12.00 Practical: Molecular Evolution (Phylogenies – PHYLIP+)
2.00-3.30 Lecture: Molecular Evolution & Comparative Genomics
3.30-5.00 Student Activity: Prepare projects

Wednesday June 3rd: Population Biology and Mapping

9.00-10.00 Lecture: Population Genetics and Gene Genealogies
10.00-11.00 Exercise: Jukes-Cantor and Rate Matrix
11.00-12.00 Lecture: Inferring Recombination Histories
2.00-3.30 Practical: DNA Sequence Analysis (PAML Phase +)
3.30-5.00 Student Activity: Prepare projects

Thursday June 4th: Integrative Genomics (IG)

9.00-10.15 Lecture: High Throughput Data, the structure of IG, $G \rightarrow F$
10.30-12.00 Practical: Statistical Alignment & Footprinting
2.00-3.30 Lecture (L): Grammars and RNA Prediction
3.30-5.00 Student Activity: Prepare projects

Friday June 5th: Integrative Genomics (IG)

9.00-10.00 Lecture: Networks and other concepts
10.00-11.00 Exercise: Stochastic Context Free Grammars
11.00-12.00 Lecture: Concepts, Data Analysis and Functional Studies
2.00-3.30 Practical: Detecting Recombinations
3.30-5.00 Student Activity: Prepare projects

Saturday June 6th: Project Discussion/Presentation

09.00-10.00 Project 1 – Proteomics
10.00-11.00 Project 2 – Inference of Pedigrees
12.00-13.00 Project 3 – RNA genes in the human genome and their evolution
13.00-14.00 Project 4 – Molecular Evolution and Annotation of Influenza

This course will run the week before “[Systems Biology Short Course](#)”. The two courses cover different topics and could both be taken with benefit.